Croatian Journal of Philosophy
Vol. XXV, No. 74, 2025
https://doi.org/10.52685/cjp.25.74.8
Received: February 11, 2024
Accepted: May 28, 2025

Words Without Intentions

J. T. M. MILLER
University of Durham, Durham, UK

A commonly held position in the literature on the metaphysics of words holds that intentions are either jointly or independently necessary or sufficient for the tokening of a word. In this paper, using a modified version of an example case created by Pavese and Radulescu (2023), I propose a counterexample in which there is no intention to token a specific word, but a word is still successfully uttered.

Keywords: Words; intentions; metaphysics of words.

A prominent claim in the literature on the metaphysics of words is that intentions are either jointly or independently necessary or sufficient for the tokening of a word (Bromberger 1989; Kaplan 1990, 2011; *inter alia*). The idea is certainly intuitively plausible. If I say the word 'plant' then this is because I have intended to token that word. I intended to utter a token of the word 'plant', and it seems natural to think that I was able to do so because of some intention to token that word.

Kaplan is perhaps most famous for holding this view. For Kaplan, intentions are necessary and sufficient for the tokening of a word: if a speaker intends to use (or token) a word, such as 'plant', then whatever output they produce just is a token of that word. As intentions are necessary for the tokening of a word, for Kaplan, it follows that if there are no intentions to utter a particular word, then that particular word cannot be uttered. As intentions are sufficient, if there are the intentions to utter a particular word, then whenever a person has the relevant intention, then the utterance is an utterance of the particular word. Hawthorne and Lepore call this the constitutive role of intentions: 'If someone intends to produce the word *w* used in a prior performance, then what comes out of their mouth (or from their pen) is a performance of *w*.' (Hawthorne and Lepore 2011: 460). For reasons that I will outline very soon, some have sought to weaken this claim from Kaplan, holding that intentions are not necessary *and* sufficient, with variations of the view holding that intentions are necessary *or* sufficient for word tokening. Let us call any view that holds that intentions are

necessary and/or sufficient for word tokening to be an 'intentionalist' account of word tokening.

As noted by Stojnić (2022: 6), both the necessity and sufficiency of intentions have been challenged, most often through one (or both) of two routes. First, we might argue that there could be two indistinguishable tokens, 'plant' and 'plant', where one of these tokens was created by a speaker (or writer) with the intention to token a word, while the other was created purely accidentally or through events that involve no human at all. It is strange to think that one of these is a word, while the other is not, especially as there are possible scenarios in which the non-intentionally created entity could still be used to express the same meaning as the intentionally created entity (see Cappelen 1999). Intentions must therefore not be necessary to tokening a word.

Second, it is common to appeal to the idea that there must be some limits of what could count as a token of a word that are independent of the intentions of the speaker or writer. A mere grunt would not count as a tokening of the word 'plant' just because I had the intention to token that word. Rather, the sound needs to be sound sufficiently like other tokens of that word to count as a token of the word. It has been suggested, for instance, that a tokening of a word must also conform to a principle of tolerance in order to count as a token of a particular word: 'Performance p is of a word w only if it meets relevant performance standards' (Hawthorne and Lepore 2011: 463). Hawthorne and Lepore therefore argue that Kaplan 'overstates the constitutive importance of intentions. In order for a performance to count as a performance of a word, it needs to fall within locally acceptable limits' (2011: 465), and that 'whether a performance counts as tolerable does not supervene on its intrinsic features, but is instead determined by certain relational facts' (2011: 463). Intentions are therefore not sufficient for the tokening of a word.

Stojnić adds a third sort of objection into the mix. Stojnić's objection is that it is 'implausible to posit a unique intention that both determines which expression is tokened and is transparent to the speaker' (2022: 5) due to considerations concerning lexical selection, and the possibility of error in lexical selection. Drawing on empirical evidence, Stojnić argues that lexical section involves the targeting of a concept, which activates a lexical item, but also activates related concepts and their associated lexical items. The most active item is then selected, but this may not be the one associated with the initially selected concept. It is this process that gives rise to selection mistakes – why we might sometimes select 'cat', even though the initially activated concept is DOG, because the concept DOG is related to the concept CAT, and its associated lexical item 'cat'. Stojnić argues that 'even if we describe lexical-selection as intentional, it is implausible to posit a unique intention that both determines which expression is tokened and is transparent to the speaker' (2022: 7). We might intend to say 'dog', but we say

'cat' because of the selection error. Instead, for Stojnić, what lexical item is selected is what is important to what word is tokened, but this has 'nothing to do with intentions' (2022: 8), and hence intentions cannot be necessary and sufficient for tokening a word.

My aim in this paper is to introduce (yet) another argument that suggests that intentions cannot be necessary or sufficient for the tokening of a word. This is not because I think there is some fundamental flaw in the existing arguments against intentional views of words. I think those arguments are good arguments (though I do not agree with the positive proposals that many of those authors then put in place instead of intentionalist accounts). Rather, I propose this because of the continued influence of intentions in the literature on the metaphysics of words and because the idea that intentions are important to tokening a word is so intuitive that even more arguments are needed to fully persuade that intentions are not a promising solution to this issue. The existing arguments focus on cases where either the sound (or ink pattern) produced fails some additional tolerance requirement or on cases where there is no intention present. My argument will introduce a case where there are some relevant intentions to utter a word and where the relevant person succeeds in uttering a word, but that even when this is the case, we cannot use intentions to determine what word is being tokened. The case that I will use is a variant of a case recently discussed by Pavese and Radulescu (2023) who themselves draw on Bratman (1984). Pavese and Radulescu use their discussion to argue against Gricean theories that analyse meaning in terms of the complex intentions of the speaker. I do not intend to make any comment on that issue here and will focus solely on the issue of tokening a word.

Imagine a society where there are two rigidly separated castes, each with its own language, each strictly unrelated to the other. The citizens speak language C, and the priests speak language P. This separation of languages is strictly enforced such that citizens and priests do not speak one another's languages. These languages are also morphologically rich (or polysynthetic). This means that in both C and P, speakers can produce highly complex word forms and single multi-morpheme forms can express what would be whole sentences in other languages. This includes, as it happens, that in both C and P, you can ask through uttering only a single word certain questions which, in English, can only be asked through uttering entire sentences composed of multiple words.

A cunning citizen devises a plan to rob a monastery where both priests and citizens live and work. As part of this plan, the cunning citizen secretly learns P, having already known C. Speaking P, he disguises himself as a priest and successfully manages to get into the monastery. After grabbing various valuable artefacts, he attempts to escape through some underground passages that lead from the monastery back into the city. In the dark, though, he gets lost. Through the

gloom, he sees someone approaching with no way to tell whether it is a priest or a citizen.

Thinking that this is his chance to get the directions he needs, and that it is too dark for anyone to know by sight whether he is a priest or a citizen, he thinks to himself: "The library and the church are close to the entrance to the underground passages. And the same sequence of sounds happens to mean two different questions in P and in C: 'Where is the access to the library?' and 'In what direction is the chapel?', respectively. If the person approaching is a citizen, they will interpret my question as being asked in C, and if a priest they will interpret it as being asked in P. As I speak both, either will give me the directions I need'.

Now, given that C and P happen to be morphologically rich in ways that English is not (but some natural languages are; see Fortescue, Mithun, and Evans 2017), both questions in C and P can be asked through uttering a single word. That is, in P it is possible to ask, 'Where is the access to the library?' through uttering only a single word, and likewise, in C, it is possible to ask, 'In what direction is the chapel?' through uttering a single word. And, as it happens, the single words in C and in P that express those questions happen to be the same phonetically. For ease of discussion, let us say that 'c' is the word in C that expresses 'In what direction is the chapel?' in English, and that 'p' is the word in P that expresses 'Where is the access to the library?' in English.

Knowing all of this, as the cunning citizen is fluent in both C and P, he decides to make the relevant sounds. If the person approaching is a priest, then they will take the cunning citizen to have uttered a word, 'p', in P, and if the person approaching is a citizen, then they will take the cunning citizen to have uttered a word, 'c', in C.

This, I argue, will be the case even though it is implausible to think that the cunning citizen had the intention to utter *either* 'c' in C or 'p' in P. If the intentionalist account of word tokening is correct, then to know what word the cunning citizen tokens, we need to know his intentions. However, it would be wrong to say that the cunning citizen intends to utter 'c' or that they intend to utter 'p'. While the cunning citizen does succeed in asking a question, then could not have intended to utter 'c' or to utter 'p' since the citizen does not intend to ask about either the library or the church specifically. His intention is not specifically to know where the church is, nor is it to know specifically where the library is, as knowing where either of these would be sufficient for our cunning citizen to make good on their escape. Under the intentionalist account of word tokening, the cunning citizen does not utter 'c' or 'p'. Indeed, under the intentionalist account, the cunning citizen does not utter any word as there is no intention to utter any specific word.

This raises a problem for intentionalist accounts of word tokening as it is also implausible to think that the cunning citizen does not token a word as whether the person approaching is a priest or a citizen, they

will interpret the cunning citizen as having uttered a (single) word. This is a case where it seems that a word is tokened, but there is no intention to token *any particular* word, in any language known by the speaker, as uttering either word would have worked to acquire the directions needed, and the cunning citizen did not know what language his addressee could speak. Any requirements proscribed by relevant tolerance principles are met in this case.

Could the intentionalist respond by arguing that the cunning citizen has an intention to utter *both* 'c' and 'p'? I argue no (for analogous reasons why this response also does not work in Pavese and Radulescu's discussion of Gricean intentions). The cunning citizen could not intend to utter both words as to intend to utter both words would involve intending to speak two languages, C and P. This would be an irrational intention as it would put the cunning citizen in danger as being able to speak both C and P is prohibited. That is, the cunning citizen does not plausibly intend to say both 'c' in C and 'p' in P because they at most intend to say one word, in whichever language would be appropriate given the unknown language that their audience speaks.

Perhaps we can reply that the issue is about being recognised as speaking both C and P, not actually speaking C and P, that puts the cunning citizen in danger. However, even if this is the case, we can here borrow two more points made by Pavese and Radulescu (2023) in response. First, let us suppose that instead of being questions about the location of the library and the chapel, the single word uttered by the cunning citizen were instead an order. Perhaps a certain order is needed to prove to the audience that the cunning citizen is a speaker of whichever language the audience speaks. Call these words 'c*' and 'p*', each a word in C* and P*. Now let 'c*' in C* be expressed in English as saying, 'Bring me only a glass of champagne' and 'p*' in P* be expressed in English as saying, 'Bring me only a glass of water'. In this case, the cunning citizen cannot intend to utter both 'c*' and 'p*' as these orders are contradictory. It is simply not possible to intend to ask for both only a glass of champagne and only a glass of water. This small change, following Pavese and Radulescu's discussion, shows that it cannot be that the cunning citizen cannot intend to utter both 'c*' and 'p*', and by extension, cannot intend to utter 'c' and 'p' in the original case.

Second, as Pavese and Radulescu also discuss in their variant of the story, it is also not plausible that the cunning citizen both particularly intends to say 'c' and particularly intends to say 'p'. Under the 'agglomeration for intentions' principle, 'if S intends to φ and S intends to ψ, then S is rationally required to have the conjunctive intention to φ and ψ' (Pavese and Radulescu 2023: 8; Bratman 1985: 220-1). Having the conjunctive intention would be irrational for the cunning citizen, and from this, agglomeration secures that he cannot have both particular intentions either. Pavese and Radulescu note that the agglomeration principle is controversial, but, also as is the case in their discussion, I

think it is the case that we also directly intuit that the cunning citizen does not have the particular intention to utter 'c', nor has the particular intention to utter 'p', meaning that the argument here (again as is also the case for Pavese and Radulescu) does not rest solely on the plausibility of the agglomeration principle. It simply is not the case that the cunning citizen only has the particular intention to utter 'c' or the particular intention to utter 'p', as to have only one of those particular intentions would be irrational in the circumstances of the thought experiment.

Another issue with the conjunctive intention response, more specific to the case of words, is that it is also unclear whether or not the cunning citizen can intend to utter both 'c' and 'p' because if this were the case, then they would token both words, and it is not clear that two words can be tokened by a single utterance. Perhaps this is possible. However, it is certainly standard in the current literature in the metaphysics of words to hold that a token is a token of one word (type). In a discussion of puns (Miller 2024), I linked this common assumption to the idea that words are like species (c.f. Wetzel 2009 in particular), and just as a particular animal can only be a token of a single species, so too a token word can only be a token of a single word (type). This last point is not a knockdown argument. However, it does suggest that for the intentionalist to respond by holding that the cunning citizen intends to, and does, utter both words would require rejecting a common assumption in the literature, present in both intentionalist and non-intentionalist views currently.

Another possible response would be to deny that the cunning citizen utters any word. A first reading of this response would amount to the claim that he only made 'mere noises' and we would explain what happens through considering the idea that the approaching stranger (wrongly) assumes that the cunning citizen has some appropriate intention to utter 'c' (or 'p'). Thus, the cunning citizen does not token any word, but communication happens because the approaching stranger assumes that they have.

This response could work but might have broader consequences. If we accept it here, then it is unclear why we should not also hold that a similar situation is occurring regularly, including even in the writing of this paper. Perhaps I do not in fact intend to utter any word when I was writing this paper, and any communicative success is due to you, the reader, assuming that I do have the intention to utter some words. This strikes me as implausible in the case of this paper, and just as implausible in the case of the cunning citizen.

Furthermore, the thought experiment assumes that the cunning citizen does intend to utter a word. The relevant point is that it is unclear that the cunning citizen intends to utter some specific word, meaning that the intentionalist does not have the means to say what word is tokened. Denying that in this case there is no word uttered would force

the intentionalist into a strange position wherein, on their view, the intention to utter 'w' is necessary and/or sufficient to token 'w', but it is *not* the case that the intention to utter any word is necessary and/or sufficient to token the type (or kind) WORD. It is unclear why we should think this is the case. Why is it that intentions are necessary and/or sufficient for the tokening of a particular word-type, but are not necessary and/or sufficient for tokening of the type WORD? I am not suggesting that this combination of views is incoherent (though see some discussion of this in Miller forthcoming a), but it certainly raises a broader point about intentionalism in that there is presumably some connection between the intention to utter a word and the intention to utter some specific word, and the case of the cunning citizen puts some pressure on what that relation actually is.

Could the intentionalist instead hold that while a cunning citizen does utter a word, they do not utter any specific word, and this is the case because while there is the intention to utter a word, there is no intention to utter a specific word? Unfortunately, this too seems to be an unattractive response. This is because it would mean that we must accept the possibility that we can utter a token of the type (or kind) WORD, without tokening some specific word-type. This is, I think, simply metaphysically impossible. Just as some object cannot have the property of being red, without being some specific shade of red, for some thing to be a word is a determinable property that must have a determinate property. Some object cannot be a word without being a token of some word-type. Pending some argument from the intentionalist that we can have a word that is not a token of some word-type, the intentionalist cannot argue that no word is uttered by the cunning citizen.

A reader here might worry that the above undermines the thought experiment that I have introduced. After all, I have above suggested that the cunning citizen does have the intention to utter a word but lacks the intention to utter a specific word. Now, I am criticising the intentionalist because I am arguing that we cannot utter a word without uttering a specific word. However, we should remember that my aim is to argue against the intentionalist account of word tokening. My own view is that the cunning citizen does utter a token of a particular word, but what I think the right view of word tokening is is irrelevant here as my aim is to argue against the intentionalist. The problem for the intentionalist, if I am right, is that we should accept that a word token is uttered, but the intentionalist cannot tell us which word the token is a token of. The question is not whether or not a word token is uttered, but that it is unclear what word that token is a token of, indicating that intentions cannot be necessary and/or sufficient for the tokening of a word.

Yet another response might be that the cunning citizen intends to utter an ambiguous word. That is, the citizen does not intend to utter

'c' in C or intend to utter 'p' in P but intends to utter 'x' which we can take to be a token of some further word. I find this response especially unconvincing. It would force the intentionalist to posit this new word, 'x', and it would be unclear which language, C or P, that 'x' is a word of. This is a significant ontological cost, and one that could lead to even further ontological costs once we allow that other speakers may also intend to speak such ambiguous words. Perhaps the words in this paper are not actually words in English because I intended to write ambiguous words, and those ambiguous words are not English words (or at least it is not clear that they are).

Furthermore, this response would allow us to create new words far too easily. Simply by intending to utter an ambiguous word, the cunning citizen has created a new word. It must be a new word as the intention is different – the cunning citizen does not intend to utter 'c' or intend to utter 'p', but intends to utter 'x', and hence, under the intentionalist view, they have uttered 'x', not 'c' or 'p'. This strikes me as clearly implausible. We cannot simply create new words in this way. Indeed, more developed forms of intentionalism (such as Stojnić 2022) build in restrictions that prevent this sort of highly personalised word creation, meaning that this response to my argument in this paper would be limited in that it would only work if the intentionalist were accepting a simpler form of intentionalism – one that Stojnić provides good reason to reject. Allowing the cunning citizen to utter a new ambiguous word would mean that creating a new word would be exceptionally easy, and hence highly ontologically costly for the view. For example, this response would mean that we could not rule out that my next utterance of 'dog' is not really an utterance of 'DOG' because I have the intention to utter a new ambiguous word. If I keep on intending that each word I utter is similarly ambiguous, we would very quickly arrive at a view that has to accept far more words than I take anyone is willing to accept.

For completeness, we can also consider another response that Pavese and Radulescu mention. Perhaps the cunning citizen has a disjunctive intention. He intends to utter *either* 'c' or 'p'. However, when considering the issue of word tokening, this response is particularly weak. The central intuition behind the intentionalist account of word tokening is that if I intend to utter a word, then my utterance is a tokening of *that* word. Words are individuated by the intentions of speakers to utter tokens of that particular word. The word 'cat', for instance, is individuated from the word 'plant' by the fact that tokens of 'cat' are the product of intentions to produce tokens of 'cat', where that is a distinct intention than the one to token the word 'plant'. It is this individuation by intention that allows intentionalist accounts to handle cases where speakers misspeak or produce tokens that do not conform to other expectations speakers have about tokens of that word. Holding that the cunning citizen has a disjunctive intention would undermine

this core claim within the intentionalist account of word tokening as we could no longer use intentions to individuate words.

This same issue also rules out a variation of this response for the intentionalist. The intentionalist might hold that the cunning citizen intends to utter either 'c' or 'p', depending on the audience. This could then be made more precise by saying that the cunning citizen has a conditional intention: to utter 'c' if the audience speaks C, and 'p' if the audience speaks P. If words are individuated by the intentions of the speaker to utter tokens of that word, and if the cunning citizen has this counterfactual intention, then it would not seem they are genuinely uttering 'c' or 'p'. Rather it would suggest they are uttering some third word which is individuated by the intention of uttering c if the audience speaks C, and p if the audience speaks P. Again, this counterfactual intention would undermine the core claim of the intentionalist account of words.

Where does this leave us? The case of the cunning citizen appears to provide a new argument against using intentions to explain how words are tokened. Unlike previous arguments against intentionalist views, the case of the cunning citizen does not deny that some intentions are present, nor does it rely on cases where our intuitions are that no word is uttered. The cunning citizen certainly intends to utter a word and is successful in uttering a word, but he does not intend to utter any particular word, and for that reason, if we accept the intentionalist account of word tokening, we cannot say what word has been tokened in this case. This is the case because, for the intentionalist, what secures that I utter a token of word w is that I intend to utter a token of word w. It is the particular intention to utter a specific word that is missing in this case, and hence the intentionalist cannot say what word is tokened. I suggest that this therefore provides another reason to reject intentionalist accounts of word tokening.

What the case of the cunning citizen does *not* do is to build a more positive view about what it is that does individuate words. Indeed, the case might even be equally problematic for some other views. For example, the case would also suggest that 'shape' cannot individuate words because if it did, we would have to hold that the cunning citizen is uttering both c and p, and also that in more mundane cases, whenever we utter a word of a particular shape we also are uttering any homonym of that word. This, I take it, is an unattractive consequence. When I say, 'I tied the bow', we do not think that I utter a token of multiple different types, all of which just happen to also be spelt (and, in some accents, pronounced) the same way.

My aim in this paper, though, is not to solve this sort of puzzle. I have argued elsewhere for my own preferred ontology (see Miller forthcoming b for a detailed discussion), but the aim of this paper is purely critical. The cunning citizen case provides a new reason to reject intentionalist accounts of word tokening.

Acknowledgements

I am grateful to the two reviewers for this journal for their comments on this paper. My thanks also to the members of the Language and Logic Research Group in Durham for many relevant discussions of intentionalist theories in the metaphysics of words.

References

Bratman, M. 1984. "Two faces of intention." *Philosophical Review* 93 (3): 375–405.

Bratman, M. 1985. "Davidson's Theory of Intention." In his *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge: Cambridge University Press, 209–224.

Bromberger, S. 2011. "What Are Words? Comments on Kaplan (1990), on Hawthorne and Lepore, and on the Issue." *Journal of Philosophy* 108 (9): 486–503.

Cappelen, H. 1999. "Intentions in words." *Noûs* 33 (1): 92–102.

Fortescue, M., Mithun, M., and Evans, N. (eds). 2017. *The Oxford Handbook of Polysynthesis*. Oxford: Oxford University Press.

Hawthorne, J. and Lepore, E. 2011. "On Words." *Journal of Philosophy* 108 (9): 447–485.

Kaplan, D. 1990. "Words." *Aristotelian Society Supplementary Volume* 64 (1): 93–119.

Kaplan, D. 2011. "Words on words." *Journal of Philosophy* 108 (9): 504–529.

Miller, J. T. M. 2024. "The Metaphysics of Puns." *Synthese* 203: 1–17.

Miller, J. T. M. Forthcoming a. "Words." In G. Dupre, R. Nefdt, and K. Stanton (eds.). *The Oxford Handbook of Philosophy of Linguistics*. Oxford: Oxford University Press.

Miller, J. T. M. Forthcoming b. *Words and Other Linguistic Entities*. Oxford: Oxford University Press.

Pavese, C. and Radulescu, A. 2023. "Meaning without Gricean Intentions." *Analysis* 83 (4): 708–716.

Stojnić, U. 2022. "Just Words: Intentions, Tolerance and Lexical Selection." *Philosophy and Phenomenological Research* 105 (1): 3–17.

Wetzel, L. 2009. *Types and tokens: on abstract objects*. Cambridge: MIT Press.